bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

BIAFLOWS:
A collaborative framework to benchmark bioimage analysis workflows

Ulysse Rubens° (ULiège), Romain Mormont° (ULiège), Volker Baecker (MRI, Biocampus Montpellier), Gino Michiels (HEPL/ULiège), Lassi Paavolainen (FIMM, HiLIFE, UHelsinki), Graeme Ball (University of Dundee), Devrim Ünay (IUE), Benjamin Pavie (VIB), Anatole Chessel (École Polytechnique), Leandro A. Scholz (Universidade Federal do Paraná), Martin Maška (Masaryk University), Renaud Hoyoux (Cytomine SCRL FS), Rémy Vandaele (ULiège), Stefan G. Stanciu (Politehnica Bucarest), Ofra Golani (Life Sciences Core Facilities, Weizmann Institute of Science, Israel), Natasa Sladoje (Uppsala University), Perrine Paul-Gilloteaux (Structure Fédérative de Recherche François Bonamy, Université de Nantes, CNRS, INSERM, Nantes, France), Raphaël Marée* (ULiège), Sébastien Tosi* (IRB Barcelona).

° These authors contributed equally to this work
* These authors jointly supervised this work (+ corresponding authors)

Abstract

Automated image analysis has become key to extract quantitative information from microscopy images, however the methods involved are now often so complex that they can no longer be unambiguously described using written protocols. We introduce BIAFLOWS, a web based framework to encapsulate, reproducibly deploy, and benchmark automated bioimage analysis workflows. BIAFLOWS helps diffusing and fairly comparing image analysis methods, hence safeguarding research based on their results by enforcing highest quality standards.

bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

As life scientists collect microscopy images of increasing size and complexity [1], the use of computational methods has become inescapable in order to extract quantitative information from these datasets. Unfortunately, most published image analysis methods are at best shared as source code, often requiring expert setup and configuration, and seldom as user friendly modules for mainstream BioImage Analysis (BIA) platforms [2][3][4]. Additionally, test images are not consistently provided with the software and it can be difficult to identify the critical parameters to adjust in order to optimize the quality of the results. Overall, this does not only limit the reusability of the methods and impedes reproducing published results [5], but it also makes it difficult to adapt these methods to process similar image datasets. To improve this situation, scientific datasets are now increasingly made publicly available through web-based software [6][7][8] and open data initiatives [9], but existing web platforms do not systematically offer advanced features such as the ability to remotely view multidimensional images, launch image analysis workflows (a software implementation of a method), and compare the results between methods or to a ground-truth reference. Having all these features integrated in the same web-based platform would offer a number of advantages and enable a fair comparison of BIA workflows on publicly available annotated image datasets (a.k.a. benchmarking). Benchmarking is known to stimulate the continuous improvements of BIA methods, simplify their deployment by specifying standard input and output data formats, and promote their wider diffusion [10]. Numerous biomedical image analysis benchmarking challenges have already been organized to compare existing methods, however all these competitions have limitations [11]: their focus is on a single bioimage analysis problem and they often rely on poorly reusable, ad-hoc, data access protocols and scripts to compute performance metrics of the workflows. Both challenge organizers and participants are therefore duplicating efforts from one challenge to the next, whereas competing workflows are rarely available in a sustainable and reusable fashion. Finally, the vast majority of annotated images released in these challenges come from medical imaging, not from research in biology: as of January 2019, only 15 out of 168 Grand-challenges datasets [12] were collected from fluorescence microscopy, arguably the most common imaging technique in the field.

Results

Within the **N**etwork of **EU**ropean **B**io**I**mage **A**nalyst**S** (NEUBIAS[1]), an important body of work focuses on channelling the efforts of bioimaging stakeholders to address these issues and ensure a better access to and assessment of existing bioimage analysis software. Together, we have envisioned and implemented BIAFLOWS (Fig. 1), *a web platform to benchmark bioimage analysis workflows on publicly shared annotated multidimensional imaging data*. Whereas some emerging bioinformatics web platforms [13][14] rely on reproducible dockerized environments and interactive Python or R notebooks to analyse scientific data accessed from a database, BIAFLOWS offers a complete framework to 1) handle multidimensional annotated images, 2) encapsulate and version BIA workflows running on most major BIA platforms, 3) remotely visualize images and workflow results and 4) assess workflows performance from widely accepted benchmark (performance) metrics. Versioned workflows can be run remotely from BIAFLOWS and are automatically benchmarked on the set of annotated images stored in projects recapitulating common image analysis problems.

[1] NEUBIAS: Network of European BioImage Analysts, COST Action CA15124. www.neubias.org

bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

All content, including images, workflows, ground-truth annotations, workflow annotations and benchmark metrics, can be browsed and interactively explored through the web interface (Fig. 1). To complement the numerical results from the benchmarking metrics, it is also possible to visualize the results of one or multiple workflows at once by synchronizing image viewers (Fig. 2). This more qualitative assessment can reveal interesting features that are not necessarily captured by the benchmark metrics (see Supplementary Methods section 7). A complete documentation (Supplementary Methods section 2) and video tutorials illustrating how to navigate the platform are available from the **Help** section of https://biaflows.neubias.org/ (user: guest, password: guest).

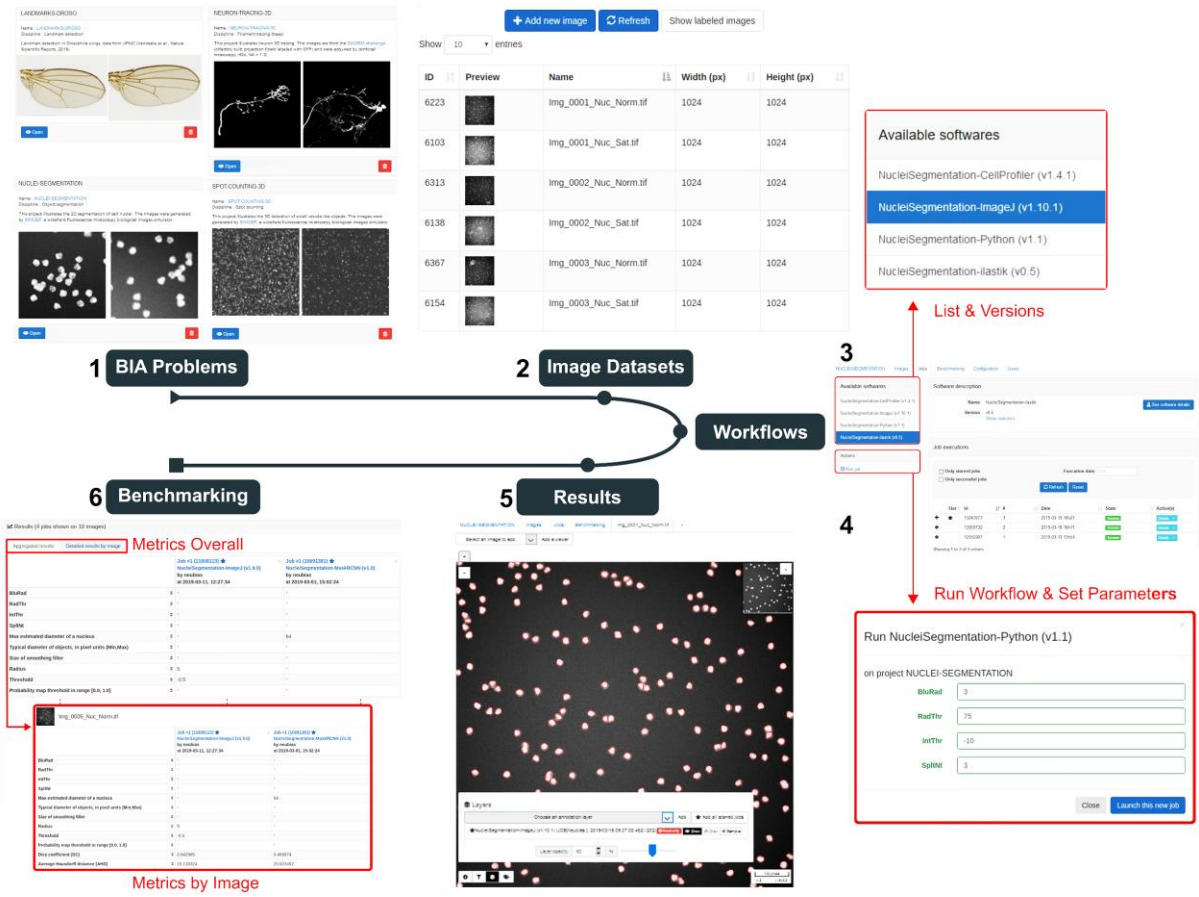


Figure 1. *The web interface of BIAFLOWS. A registered user first selects a bioimage analysis problem from the set of available classes (1) and browses compatible datasets (2) from this project. A workflow to process these images can then be selected from a list (3), and associated parameters adjusted (4). Finally the results of the workflow can be visualized as an overlay layer in the remote image viewer (5), and the associated benchmark metrics are reported (6) for all images of the project.*

bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

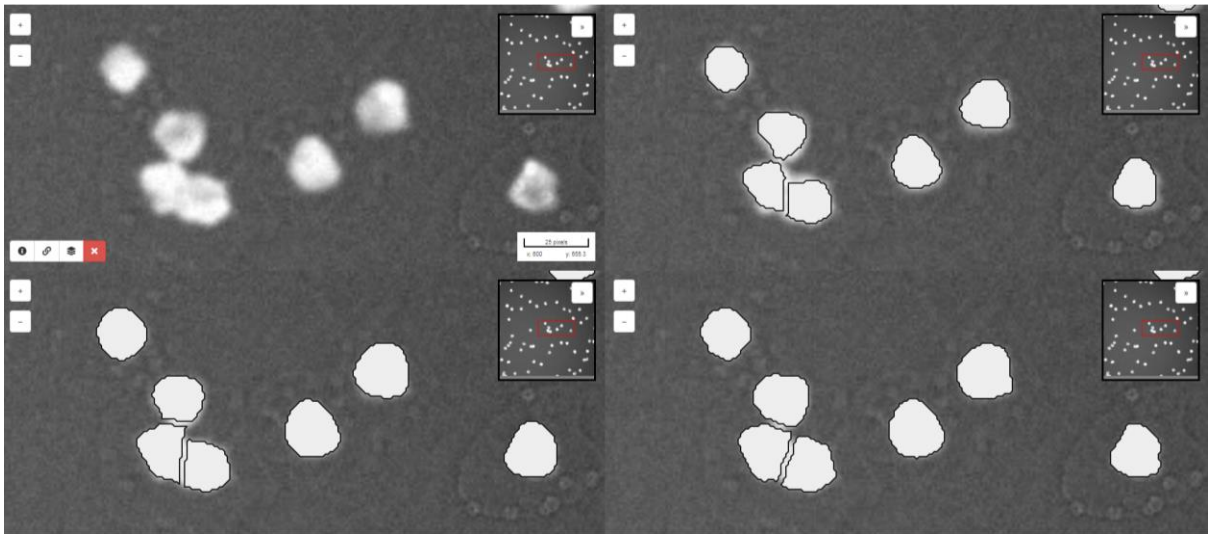


Figure 2 BIAFLOWS enables to synchronize several image viewers and to display results from different workflows as overlays. This can be useful to compare several workflows and grasp qualitative differences in their results. On this figure, results for nuclei segmentation are shown for ImageJ (upper right), Python (lower left) and CellProfiler (lower right) workflows integrated to the system for this bioimage analysis problem. The original image is also displayed without workflow results overlay (upper left).

BIAFLOWS is based on Cytomine [15], an open-source web platform developed for the collaborative annotation of high-resolution biomedical images (especially large 2D histology images). The main features that were developed for BIAFLOWS to enable the benchmarking of bioimage analysis workflows are described in this section. First, a module has been added to support the upload of microscopy multidimensional images (C, Z, T) in OME-TIFF format, as well as their remote visualization in a viewer enabling to navigate through image slices, adjust contrast and efficiently toggle annotation layers. Next, the core architecture has been completely re-designed to support the remote and remote execution of bioimage analysis workflows by encapsulating the workflows and their original software environment in Docker images, and describing their interface by an extended version of Boutiques, a rich application description schema [16]. BIAFLOWS is designed to monitor trusted user spaces hosting workflow images and automatically pull new or updated workflows (Fig. 3, DockerHub). In turn, workflow Docker images are automatically built whenever a new release is triggered from their associated source code repository (Fig. 3, GitHub). To ensure reproducibility, all workflow images are versioned and permanently accessible from the system, moreover they can be launched on virtually any type of computational resources (including high-performance computing and multiple server architectures). This is seamlessly achieved in BIAFLOWS by converting the workflow images to a compatible format (Singularity, [17]) and dispatching them to the target computational resources using a workload manager (SLURM, [18], Fig. 3, additional computing servers).

bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

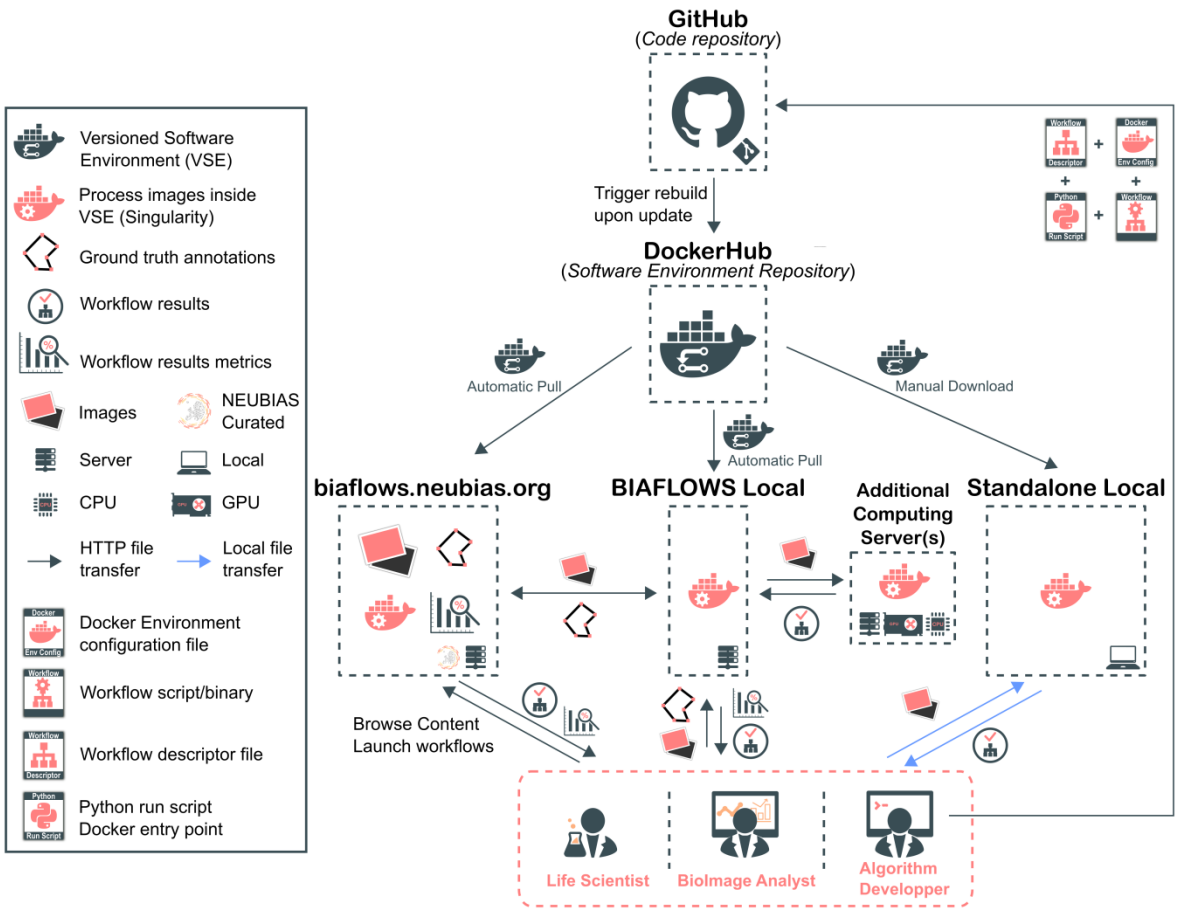


Figure 3. *BIAFLOWS architecture and application scenarios. Workflows are stored in a trusted code repository (GitHub). Workflow images are automatically built and available through DockerHub upon code update. An instance of BIAFLOWS monitoring DockerHub automatically pulls new or updated workflows. This can either be NEUBIAS benchmarking online instance or a local instance installed by a user to manage and process personal data. Workflow images can also be manually downloaded and used to process a local folder of images.*

To demonstrate the possibilities of our framework, an online instance of BIAFLOWS is available from this URL: https://biaflows.neubias.org/ (Fig. 3, BIAFLOWS NEUBIAS). It is populated with some selected image datasets and associated workflows and can be browsed in read only mode from the guest account (user: guest, password: guest). Currently 14 annotated datasets and 16 workflows are available and they correspond to 8 classes of common BIA problems: object detection/counting, object segmentation, pixel classification, particle tracking, object tracking, filament tree tracing, filament network tracing, and landmark detection (Fig. 4 and Fig. 5, as well as Supplementary Methods section 1, Table 1). To enable the automatic execution and benchmarking of the workflows, some standard object annotations formats were specified for each class of problem (see Supplementary Methods, section 5). The code to compute benchmark metrics was adapted from challenges [12] and scientific publications [20]. The annotated datasets were imported from existing challenges (see section “Methods”) or created from synthetic image generators [19]. To showcase the versatility of the platform, the workflows consist of a mixture of standalone

bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

software, scripts and pipelines (ImageJ/Fiji, ICY, CellProfiler, Vaa3D, ilastik and Python). Some workflows leverage deep learning Python libraries (Keras, PyTorch). In addition, to enhance their visibility, all workflows hosted in the system are referenced from Bioimage Informatics Search Index (http://biii.eu/), an online repository of bioimage analysis tools maintained by NEUBIAS. Interested developer can package their workflows and make them available for benchmarking from the online instance (see Supplementary Methods, section 4) by sending a request to BIAFLOWS administrators[2]. For full flexibility, BIAFLOWS can also be deployed and populated locally (images and workflows) so as to be used as a local image management and analysis solution (Fig. 3 BIAFLOWS local, see Supplementary Methods section 3). To simplify this process, migration tools were developed to transfer content between existing BIAFLOWS instances (see Supplementary Methods, section 6). Finally, for maximal flexibility, all content can be accessed programmatically through a RESTful application program interface (see Supplementary Methods, section 6) and BIAFLOWS workflows can also process local images independently of any BIAFLOW server too (Fig. 3, standalone local, Supplementary Methods section 6).

Discussion

End users running a bioimage analysis workflow from a specific tool (e.g. ImageJ) on their own computer might have very hard time assessing the quality of the results, knowing whether the parameters are well adjusted for their own images, or finding whether there would be a better achieving workflow to tackle the same problem. Hence, it is critical to have an open source, web-based, framework where several *parameter optimized* workflows from any tool can be run, and where their results can be visualized conjointly and objectively compared to a ground truth reference. Algorithm developers also benefit from using common annotated datasets to compare their workflows to other existing solutions, and hence complement the performance they could assess on project specific data. Advertising the performance of bioimage analysis workflows through a public portal is probably the best way to increase their wide diffusion.

BIAFLOWS is a web-based platform enabling the testing and benchmarking of automated bioimage analysis workflows producing object annotations from raw images, a critical step when extracting quantitative information from scientific images. The online instance managed by NEUBIAS is available at https://biaflows.neubias.org and is populated with a starting set of annotated bioimage datasets and associated workflows corresponding to common bioimage analysis problems; these workflows can be run and benchmarked from a simple web browser. BIAFLOWS provides a framework to integrate new workflows and it is fairly simple to add new content (images, annotations, workflows) or to migrate existing content between different BIAFLOWS instances. We demonstrate BIAFLOWS use cases for 15 annotated datasets and 15 workflows spanning over 8 classes of bioimage analysis problems. This clearly only represents a subset of all existing problems but the problems included are major building blocks of many other, more complex, tasks (e.g. object based colocalization analysis). Additionally, new problem classes can easily be included. For example, we plan to include blinking events detection in the context of super-resolution

[2] Send a request to biaflows@neubias.org

bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

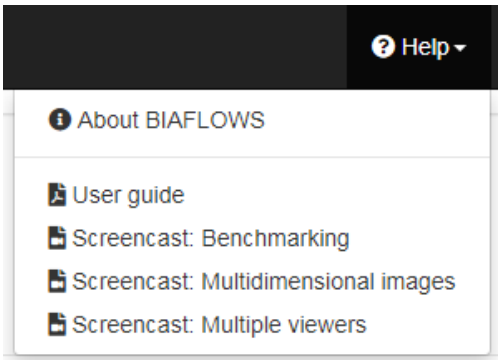
localization microscopy and landmark matching for image registration. To increase the content of BIAFLOWS, calls for contribution will now be launched within NEUBIAS to gather more microscopy annotated images and incentive developers to compare their methods to other existing methods already integrated to the platform. BIAFLOWS is designed to expose the functional parameters of a bioimage analysis workflow, and to provide default *optimal* values (Fig. 1, step 4). The *optimality* is left to the responsibility of the developer integrating a workflow to the platform. Every workflow hence comes pre-configured with a set of default parameters to guarantee optimal results for the set of images provided of a given project. These parameters can however be freely modified by the user to assess their impact on the final results. As an added value, BIAFLOWS integrated bioimage analysis workflows can be reproducibly deployed on most software platform. This can be achieved within a local instance of BIAFLOWS, or without BIAFLOWS on any machine or HPC environment running Docker. This feature is critical, especially for workflows requiring a complex deployment (e.g. compilation with multiple software dependencies), but also when deploying workflows on home platforms so as to avoid non reproducibility related to software version (e.g. JAVA, home platform, installed plugins). Interestingly, BIAFLOWS could also be used to trigger workflow automated runs spanning part of their functional parameter space. This could be helpful to automatically optimize these parameters for the set of images of a project, or to assess the impact of parameter misadjustements (a critical feature from usability point of view). Overall, BIAFLOWS addresses a number of critical requirements currently posed to open image analysis for life sciences, including (i) sharing versioned image analysis workflows in a reproducible way, (ii) exposing the functional parameters of the workflows, (iii) providing a standardized way to access results, (iv) providing compatible annotated images illustrative of biology research projects, and (v) computing relevant performance metrics to compare workflows on these images.

bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Methods

Accessing BIAFLOWS online instance

BIAFLOWS online instance can be accessed at https://biaflows.neubias.org and browsed in read-only mode from the guest account (username: *guest* password: *guest*). The platform is maintained by NEUBIAS (http://neubias.org). Video tutorials illustrating how to navigate and use the platform are available from the **Help** section of the website.



A list of all the content currently available from the website is provided (Supplementary material section 1, Table 1) and some problems are illustrated in Fig. 4 and Fig. 5. The image datasets have been selected to recapitulate common BIA analysis tasks: spot detection (2D/3D), nuclei segmentation (2D/3D), nuclei tracking (2D), landmark detection (2D), filament tracing (3D), and tissue detection (2D) in whole-slide histology images. All image datasets are imported from previously organized challenges (DIADEM [21], ISBI Cell Tracking Challenge [22], ISBI Particle Tracking Challenge [23], Kaggle Data Science Bowl 2018 [24]), created from synthetic data generators (CytoPacq [25], TREES toolbox [26], Vascusynth [27], SIMCEP [28]), or contributed by NEUBIAS members [37]. To showcase the versatility of the platform, the image analysis workflows available to process these images are running on different BIA platforms: ImageJ macros [29], Icy protocols [30], CellProfiler pipelines [31], Vaa3D plugins [32], ilastik pipelines [33], Python scripts leveraging Scikit-learn [34] for supervised learning algorithms, and Keras/PyTorch [35][36] for deep learning. These workflows were contributed by members of NEUBIAS work group 5, or imported from existing challenges.

bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

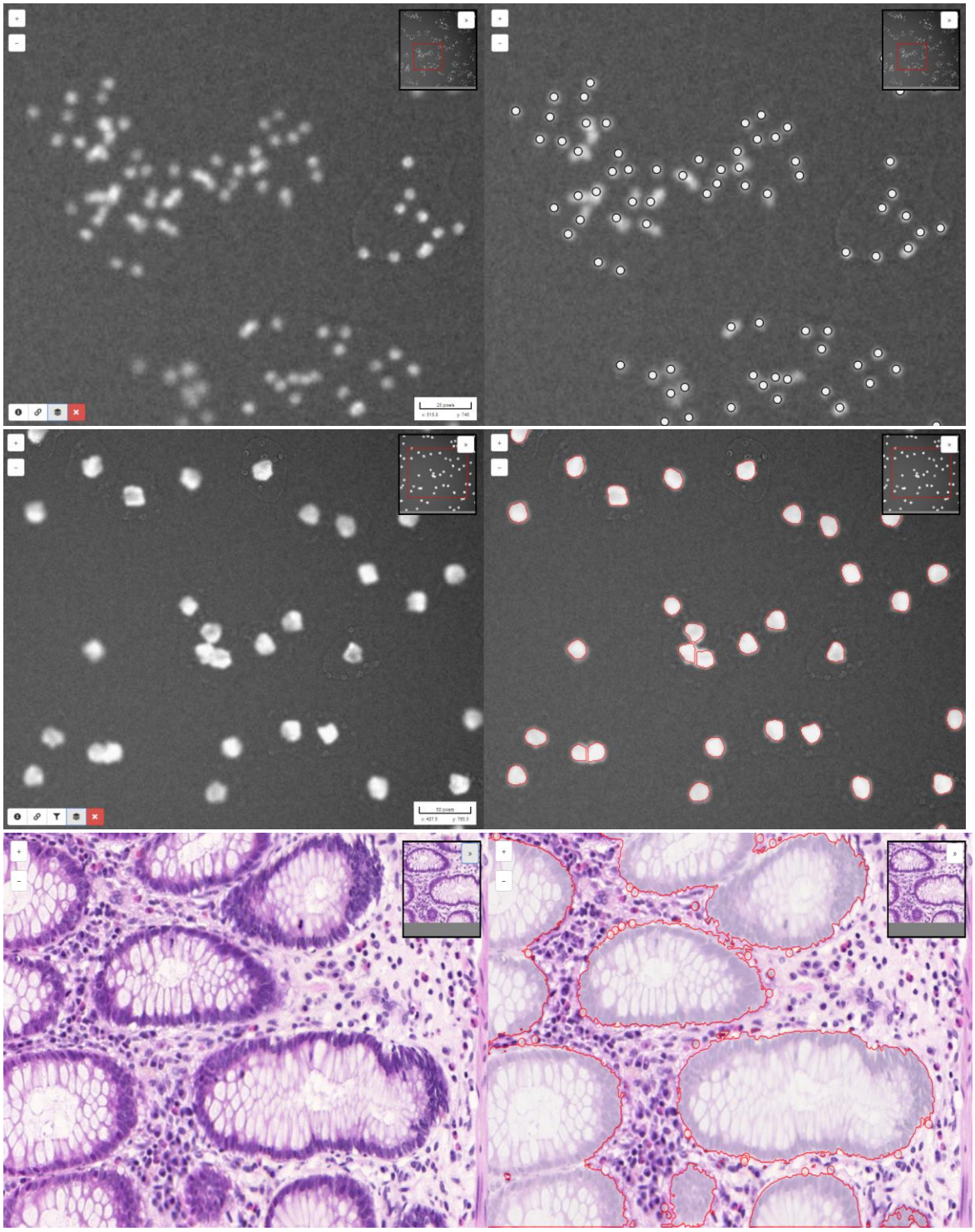


Figure 4 Some images illustrating BIA problems and associated workflow outputs from BIAFLOWS online instance. Original image (left) and workflow results overlay (right). 1. Spot / object detection & counting, synthetic image displaying spots generated by SIMCEP [38]. 2. Object segmentation, synthetic image displaying nuclei generated by SIMCEP [38]. 3. Pixel classification, red areas circle pixels classified as gland, image from 2015 MICCAI challenge of gland segmentation [38].

bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

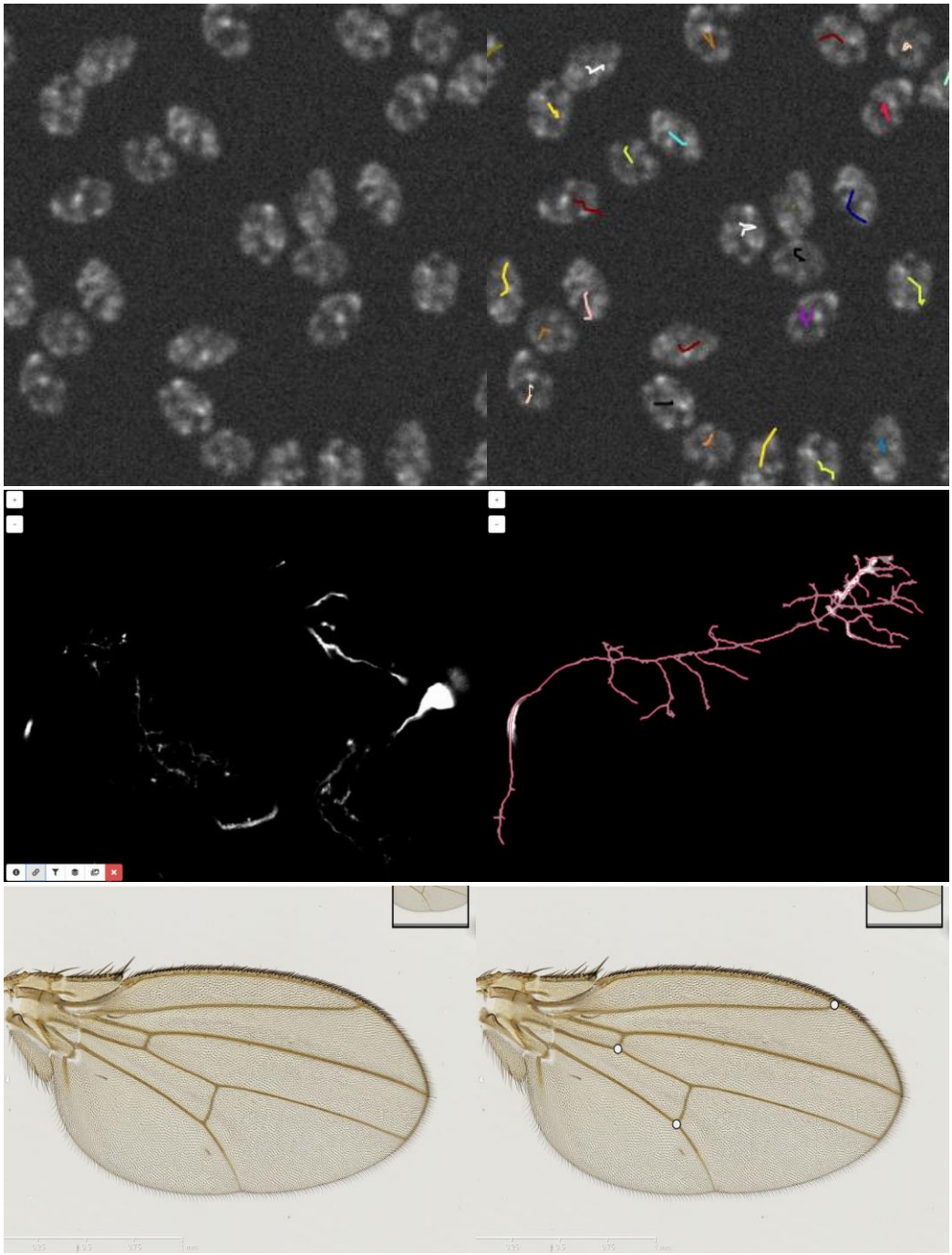


Figure 5 Some images illustrating BIA problems and associated workflow outputs from BIAFLOWS online instance. Original image (left) and workflow results overlay (right). 1. Particle / object tracking, synthetic time-lapse displaying non-dividing nuclei generated by CytoPACQ [25], single frame + dragon tail tracks showing nuclei past positions. 2. Neuron tracing, Z-stack from DIADEM challenge [21], single slice + dilated Z-projection of traced skeleton (red). 3. Landmark detection, Drosophila wing, image from UPMC [37].

bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Acknowledgments

This project is funded by COST CA15124 (NEUBIAS). BIAFLOWS is developed by NEUBIAS (http://neubias.org) work group 5 and it would not have been possible without the great support from NEUBIAS vibrant community of bioimage analysts, and the dedication of Julien Colombelli and Kota Miura. Local organizers of the hackathons who have fostered the development of BIAFLOWS are also greatly acknowledged: Chong Zhang, Gabriel Martins, Julia Fernandez, Peter Horvath, Bertrand Vernay, Aymeric Fouquier d'Herouel, Andreas Girod. **UR** and **RV** were supported by ADRIC Pôle Mecatech Wallonia grant. **RaM** was supported by IDEES grant with the help of the Wallonia and the European Regional Development Fund (ERDF). **LP** was supported by Academy of Finland (grant 310552). **MM** was supported by the Czech Ministry of Education, Youth and Sports (project LTC17016). **SGS** acknowledges the financial support of UEFISCDI grant PN-III-P1-1.1-TE-2016-2147 (CORIMAG). **VB** and **PPG** acknowledge the France-BioImaging infrastructure supported by the French National Research Agency (ANR-10-INBS-04).

Author contribution

ST and **RaM** conceptualized BIAFLOWS, supervised its implementation, contributed to all technical tasks and wrote the manuscript. **UR** worked on the core implementation of BIAFLOWS with contributions from **GM** and **RH**. **RoM** implemented several modules to interface bioimage analysis workflows and the content of the system. **ST, MM** and **DU** implemented the module to compute benchmark metrics. **VB**, **RoM**, **BP**, **MM, RV, LAS** and **LP** integrated workflows and tested the system. **AC**, **DU**, **OG,** and **GB** organized and collected content (image datasets, simulation tools). **SGS**, **NS** and **PG** provided extensive feedback on BIAFLOWS and contributed to the manuscript. All authors took part in reviewing and improving the manuscript.

Competing interests

RaM and **RH** are co-founders of the non-profit cooperative company Cytomine SCRL FS.

Data availability

All images and annotations can be downloaded from the online instance of BAIFLOWS at https://biaflows.neubias.org/.

Code availability

BIAFLOWS is an open source project and its code source can be downloaded at https://github.com/Neubias-WG5. The procedure to install a local instance of BIAFLOWS is described in Supplementary material (section 3).

bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

Additional information

Methods to access BIAFLOWS online instance including a complete documentation, and video tutorials, as well as the procedure to install and populate a local instance of BIAFLOWS (images, annotations, workflows) and migrate content between instances are available in supplementary information.

References

1. Ouyang, W., & Zimmer, C. (2017). The imaging tsunami: computational opportunities and challenges. Current Opinion in Systems Biology, 4, 105-113.
2. Kevin W Eliceiri et al. (2012). Biological imaging software tools. Nature Methods volume 9, pages 697–710.
3. Carpenter, Anne E., Kamentsky, Lee and Eliceiri, Kevin W. (2012). A call for bioimaging software usability. Nature methods 9, no. 7: 666-670.
4. Schneider, C. A.; Rasband, W. S. & Eliceiri, K. W. (2012), NIH Image to ImageJ: 25 years of image analysis, Nature methods 9(7): 671-675.
5. Munafò, M. R. et al. (2017). A manifesto for reproducible science. Nat. Hum. Behav. 1, 0021.
6. Ellenberg et al. (2018). A call for public archives for biological image data. Nature Methods October, Pages 849–854.
7. Chris Allan, et al. (2012) OMERO: flexible, model-driven data management for experimental biology. Nature Methods 9, 245–253.
8. Kristian Kvilekval, et al. (2010). Bisque: a platform for bioimage analysis and management, Bioinformatics, Volume 26, Issue 4, 15 February, Pages 544–552.
9. Eleanor Williams, et al. (2017). Image Data Resource: a bioimage data integration and publication platform. Nature Methods volume 14, pages 775–781.
10. Vandewalle P. (2012). Code sharing is associated with research impact in image processing. Computing 799 in Science & Engineering, 14(4):42–47.
11. Maier-Hein L. (2018). Why rankings of biomedical image analysis competitions should be interpreted with care. Nature Communications, 9:5217.
12. https://grand-challenge.org/
13. Perkel J.M. (2018). A toolkit for data transparency takes shape. Nature Technology Feature. August, 560,513-515.
14. Grüning BA, Rasche E, Rebolledo-jaramillo B, et al. (2017). Jupyter and Galaxy: Easing entry barriers into complex data analyses for biomedical researchers. PLoS Comput Biol.;13(5):e1005425.
15. Marée et al. (2016). Collaborative analysis of multi-gigapixel imaging data with Cytomine, Bioinformatics, 32(9):1395-401.
16. Tristan Glatard et al. (2018). Boutiques: a flexible framework to integrate command-line applications in computing platforms, *GigaScience*, Volume 7:5, 1 May.
17. Kurtzer GM, Sochat V, Bauer MW. (2017). Singularity: Scientific containers for mobility of compute. PLoS ONE 12(5): e0177459.

bioRxiv preprint doi: https://doi.org/10.1101/707489; this version posted July 19, 2019. The copyright holder for this preprint (which was not certified by peer review) is the author/funder. All rights reserved. No reuse allowed without permission.

18. Yoo A., Jette M., and Grondona M. (2003). Slurm: Simple Linux Utility for Resource Management, Job Scheduling Strategies for Parallel Processing, volume 2862 of Lecture Notes in Computer Science, pages 44-60, Springer-Verlag.
19. Svoboda D, Kozubek M, Stejskal S. (2009). Generation of digital phantoms of cell nuclei and simulation of image formation in 3D image cytometry. Cytometry A.; 75: 494–509.
20. Kozubek M. Challenges and Benchmarks in Bioimage Analysis. (2016). Adv Anat Embryol Cell Biol. 219:231-62.
21. http://diademchallenge.org/
22. http://celltrackingchallenge.net/
23. http://bioimageanalysis.org/track/
24. https://www.kaggle.com/c/data-science-bowl-2018
25. https://cbia.fi.muni.cz/simulator/
26. https://www.treestoolbox.org/
27. http://vascusynth.cs.sfu.ca
28. http://www.cs.tut.fi/sgn/csb/simcep/tool.html
29. https://imagej.nih.gov/ij/developer/macro/macros.html
30. http://icy.bioimageanalysis.org/plugin/Protocols
31. https://cellprofiler.org/examples/
32. https://github.com/Vaa3D
33. https://www.ilastik.org/
34. https://scikit-learn.org
35. https://keras.io/
36. https://pytorch.org/
37. Vandaele R. et al. (2018). Landmark detection in 2D bioimages for geometric morphometrics: a multi-resolution tree-based approach, Scientific Reports 8, Article number; 538.
38. https://warwick.ac.uk/fac/sci/dcs/research/tia/glascontest/